



IFWO

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/10/775,501A

TIME: 11:47:24

Input Set : A:\84500-1.app

Output Set: N:\CRF4\09132004\J775501A.raw

3 <110> APPLICANT: Peltonen, Leena
 4 Enattah, Nabil
 5 Jaervelae, Irma
 6 Sahi, Timo
 7 Savilahti, Erkki
 8 Terwilliger, Joseph
 9 National Public Health Institute
 11 <120> TITLE OF INVENTION: Identification of a DNA Variant Associated With Adult
 12 Type Hypolactasia
 14 <130> FILE REFERENCE: 084500-000100US
 16 <140> CURRENT APPLICATION NUMBER: US 10/775,501A
 17 <141> CURRENT FILING DATE: 2004-02-09
 19 <150> PRIOR APPLICATION NUMBER: EP 01119377.8
 20 <151> PRIOR FILING DATE: 2001-08-10
 22 <150> PRIOR APPLICATION NUMBER: EP 01119528.6
 23 <151> PRIOR FILING DATE: 2001-08-14
 25 <150> PRIOR APPLICATION NUMBER: US 60/315,955
 26 <151> PRIOR FILING DATE: 2001-08-31
 28 <150> PRIOR APPLICATION NUMBER: WO PCT/EP02/08963
 29 <151> PRIOR FILING DATE: 2002-08-09
 31 <160> NUMBER OF SEQ ID NOS: 17
 33 <170> SOFTWARE: PatentIn Ver. 2.1
 35 <210> SEQ ID NO: 1
 36 <211> LENGTH: 180
 37 <212> TYPE: DNA
 38 <213> ORGANISM: Homo sapiens
 40 <220> FEATURE:
 41 <223> OTHER INFORMATION: lactase persistence type intron 13 of the MCM6 gene with
 42 single nucleotide polymorphism (SNP) t substituted by c at
 43 position -13910 5' from the intestinal lactase-phlorizine
 44 hydrolase (LPH) gene
 46 <400> SEQUENCE: 1
 47 accttttcatt caggaaaaat gtacttagac cctacaatgt actagtaggc ctctgcgctg 60
 48 gcaatacaga taagataatg tagcccttg cctcaaagga actctctcc ttaggttgca 120
 49 tttgtataat gtttgatttt tagattgttc tttgagccct gcattccacg aggataggtc 180
 52 <210> SEQ ID NO: 2
 53 <211> LENGTH: 180
 54 <212> TYPE: DNA
 55 <213> ORGANISM: Homo sapiens
 57 <220> FEATURE:
 58 <223> OTHER INFORMATION: lactase persistence type intron 9 of the MCM6 gene with
 59 single nucleotide polymorphism (SNP) a substituted by g at
 60 position -22018 5' from the intestinal lactase-phlorizine

(pg.6)

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/10/775,501A

TIME: 11:47:24

Input Set : A:\84500-1.app

Output Set: N:\CRF4\09132004\J775501A.raw

```

61      hydrolase (LPH) gene
62      <400> SEQUENCE: 2
63      taagaacatt ttacactctt cagtataaag aagtcagaat acccctaccc tatcagtaaa 60
64      ggctataag ttaccattaa aaagatgtcc ttaaaaaacag cattctcagc tgggcgcggt 120
65      ggctcacacc tttgtcccag tactttggga agccgaggtg ggtggatcac ctgaggtcag 180
66      <210> SEQ ID NO: 3
67      <211> LENGTH: 3213
68      <212> TYPE: DNA
69      <213> ORGANISM: Homo sapiens
70      <220> FEATURE:
71      <223> OTHER INFORMATION: lactase persistence type intron 13 of the MCM6 gene with
72      single nucleotide polymorphism (SNP) t at position -13910
73      5' from the intestinal lactase-phlorizine hydrolase (LPH)
74      gene
75      <400> SEQUENCE: 3
76      atcagagtca ctttgatatg atgagagcag agataaacag atttgttgca tgtttttaat 60
77      ctttggtatg ggacatacta gaattcactg caaatacatt tttatgtaac tgttgaatgc 120
78      tcatacgacc atggaattct tccctttaaa gagcttggtg agcatttgag tgtagttggt 180
79      agacggagac gatcacgtca tagtttatag agtgcataaa gacgtaagtt accatttaat 240
80      acctttcatt caggaaaaat gtacttagac cctacaatgt actagtaggc ctctgcgctg 300
81      gcaatacaga taagataatg tagtccctgg cctcaaagga actctctctc ttaggttgca 360
82      tttgtataat gtttgatttt tagattgttc tttgagccct gcattccacg aggataggtc 420
83      agtgggtatt aacgaggtaa aaggggagta gtacgaaagg gcattcaagc gtcccatctt 480
84      cgcttcaacc aaagcagccc tgcgttttcc tagttttatt aataggtttg atgtaaggtc 540
85      gtctttgaaa aggggggttg gctttttttt acagtgtgac tgagggtataa tttataaaaa 600
86      gggaaatgta tggcatggtg agttttttca catacatcct tgtgaatacc cagctcaaga 660
87      tccaaaacat ttccataatt tcagaaagtt ccaaaccctt gcctcttttc agtcttagcc 720
88      ctcttccctt gaagtaacca ctgttcogac ttcaatcact acttttatcc cacagggtta 780
89      ttttttggct tttttccact aaattttcaa attctttgat atggtacttt actattgacg 840
90      aagtactttc acactagggt atttaatat ctttgattca cccaatattt agggaaacacc 900
91      tgtaggggac aaaaaatgaa tgagagcccc tgccttccat tgcgtgtaat ctgggtggaa 960
92      cgagacatgt atttaattaa gcatgtaaaa aatagagtgg gtgatgaaat aatctatata 1020
93      ctaaatcccc atgacacaca gtttacctat gtaacaaacc tgcattgtga ccccggaacc 1080
94      taaaatataa gttggaaatt aaaaaaaaaa gagagggaga atagagcatc acaaccagag 1140
95      tgctgagatg aattacttta ttaccaaaga aggaggagga ctgaggaggg tgccgacgtt 1200
96      taaacccagt cactgaaggg tgtgcagaat ttggataggc aagataccct gggacaaggt 1260
97      cattctaaaa ccatgctaac atttgtactt ttttttcat tgtgatagtt cctgaaatga 1320
98      gttgcataaa actggtacat gtcttagggc agtctctaatt tgatttttat tttgttctat 1380
99      ttttaaaaat tagtcttcaa atagcagatt cacatgatat taaaatatat gcacataaat 1440
100     tatatacaca aatatatttt ctgaatgaaa tttagtatct gcatatattt aagagctatt 1500
101     tctgtctcat atgttcataa tcttcaccca ttaaaaaaac ttttgtagg cttttctcac 1560
102     tetaagatta taaaaaatcc tccattatt tacctagcta gttttctagt tgttccaaaa 1620
103     ccatttattg aacaatccat ctttttgaca ctgggttggc atgccttaat tatatattct 1680
104     tgtgtgtggt aggatctcct tttggacttt ccattctggt cattgagtct tatcagctcc 1740
105     tcttacattg gtaccatgat gttttaatct atggggcttt gtagtttaaa tgtagggtca 1800
106     gttccagcgc attgttctct atcagctggt aggaacttag aaatcagctt gctctgtttt 1860
107     aaagaaaaac ctggtatttt tttatcagta taacattcta tttatattaa cttgaagaat 1920
108     tgaaaacatc tatgattttt cctattcagt aacgtatcac ttagaatagg ttaggttgta 1980
109     ctactataaa atctcagctg cataaaacaa tttttttttg cttgtgctac acatccatta 2040

```

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/10/775,501A

TIME: 11:47:24

Input Set : A:\84500-1.app

Output Set: N:\CRF4\09132004\J775501A.raw

```

115 ggtcatcaag ggactcacct tgtcaagtta ctcagagatt caggctgata taaagggttg 2100
116 atcttgacat acgctttcat gatgacagaa agcaggggaag agaagggtgt gagccatgtg 2160
117 ctttctcccc cttctatcca gaaatgacac atactcacat ttcattcgcc agagaaatta 2220
118 acatggcccc tcctaagttc aaatggatag agaaatgcct tcctaccagg tgcccagaat 2280
119 tagaagagca aacattttgtg aacagttctg agtaccacaa ataccgttat ctttccactt 2340
120 aagtcttctg tttcactcag tagtgcttta aacttttctt catatgtttt tcagtgtttc 2400
121 ttgttgaaat tcttgatatt ttatcatgtt tgttcgtact gggagtagcc tttttttcca 2460
122 tttcattttc tggctggttt cattgctggt tgtttttttg ttttgttttg tttttgagat 2520
123 ggagtctcac tctgtcgccc aggctggagt gcagtgtcac aatctcggt cactgcaacc 2580
124 tctgcctccc aggttcaagc gattcttctt tctcagcctc ctgagtagct gggattacag 2640
125 gcatgtgcca ccatgcccag ctaatttttt atatttttag tagagatggg gtttctccat 2700
126 gttggtcagg ctggtctcaa actcccaatc tcagggtgat cgctgcctc tgccttccaa 2760
127 agtgctggga ttatagacat gagccaccgt gcctggccta gttcttatgg gatgtatatg 2820
128 tctttggatt catatgatat gtatatatgt ttatatctt acaagtacat acctaggagt 2880
129 ggaattgttg ggtcataggt taatgcatgt ttttctgcca aacagttgtg tcaattttctg 2940
130 ttttcaccgc tgtgaatgag agttgttcta ccttcttgac aacacttgat attgtcagtc 3000
131 attttagcca ttctggtgaa tttatagtgc tatttctgtg tgtgtaagag agagaatgag 3060
132 agagggtgtt tgtgagaaaa ccaaagcaac actgtgagag tgtgtgtgtt tgtgagaaaa 3120
133 ccaaaataca tactactgtg atttcattgg gagaaaatct gtttggtata tcaaaaaaag 3180
134 tagcttaatt acttcatcat tattggttta ggt 3213

```

137 <210> SEQ ID NO: 4

138 <211> LENGTH: 1296

139 <212> TYPE: DNA

140 <213> ORGANISM: Homo sapiens

142 <220> FEATURE:

143 <223> OTHER INFORMATION: lactase persistence type intron 9 of the MCM6 gene with
 144 single nucleotide polymorphism (SNP) a at position -22018
 145 5' from the intestinal lactase-phlorizine hydrolase (LPH)
 146 gene

148 <400> SEQUENCE: 4

```

149 taagaacatt ttacactctt cagtataaag aagtcagaat acccctaccc tatcagtaaa 60
150 ggccatataag ttaccattaa aaagatgtcc ttaaaaaacag cattctcagc tgggcacggg 120
151 ggctcacacc tttgtcccag tactttggga agccgagggt ggtggatcac ctgaggtcag 180
152 gagttcgaga ccagcctggc caacatggcg aaaaccatt ttcttacta aaaatacaaa 240
153 aattagccgg gcatggtggc ggggtgctgt ggtcccagct actcaagagg ctgaggtggg 300
154 aggatcactg agcccaggag gtggaggctg cattgagcca agattgtgcc actgcactcc 360
155 agcctgggtg acagagcgag actctgtctc aaaaaaacca aaacaaaaaa aaccagcat 420
156 tctttagtaa ataattcata gttttcttca tctagaattt aaaattgtga tagttgatca 480
157 gcatgtcctg agcacgtgtg tttgctgtta ctagtttaga tcggtagatg tgtatataag 540
158 ttataggtat aaaatcaatc ctgagttgac acaaggtttt gatgttgagt acaagtacag 600
159 taagtgtata tttttagtta tgctcttagt ttttaagtcaa ttgtgtggtt ctttctagct 660
160 ttaggatctg ttgaattatc ttccttagaa aagggagtta agaattctca cttacctatc 720
161 ttctacttgt ttggagaata gaagagtccc tgtggtagca gactttgtga gtttacttgt 780
162 aattttccat ctgaaagact gttcttggtt ttcgtgatga agtcttgctc tgtcgcccag 840
163 gctggagtgc agtggtgcaa ccttggtcca ctgcaacctc tgcctcccgg gttcaagcaa 900
164 ttctcctgcc tcagcctccc gagtatctgg gattacaggt gcacaccacc acacctggct 960
165 aatttttgta ttttcagtag agacgggggt tcaccatgtt ggccaggctg gtctcgaact 1020
166 cttgacctca tgatcagccc acctcagcct tccaaagtgc tgggattaca ggtgtgagcc 1080
167 cccacactcg gccgttggtg ttttttaaga gacagggctc cactctgtca cctaacctgg 1140

```

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/10/775,501A

TIME: 11:47:24

Input Set : A:\84500-1.app

Output Set: N:\CRF4\09132004\J775501A.raw

```

168 agtacagtgg caatcatggc tcaactgtaac ctcaaatgcc cggccttagt gaagegttct 1200
169 tcctgccttg gcctcccaaa gtgctgggat tacaagtgtg agccatgcat ccagcttgaa 1260
170 agacagcttc ttaggcttga tttgtttggt tacagg                               1296
173 <210> SEQ ID NO: 5
174 <211> LENGTH: 3213
175 <212> TYPE: DNA
176 <213> ORGANISM: Homo sapiens
178 <220> FEATURE:
179 <223> OTHER INFORMATION: lactase persistence type intron 13 of the MCM6 gene with
180     single nucleotide polymorphism (SNP) t substituted by c at
181     position -13910 5' from the intestinal lactase-phlorizine
182     hydrolase (LPH) gene
184 <400> SEQUENCE: 5
185 atcagagtca ctttgatatg atgagagcag agataaacag atttgttgca tgtttttaat 60
186 ctttggtatg ggacatacta gaattcactg caaatacatt tttatgtaac tgttgaatgc 120
187 tcatacgacc atggaattct tccctttaaa gagcttggtg agcatttgag tgtagttggt 180
188 agacggagac gatcacgtca tagtttatag agtgcataaa gacgtaagtt accatttaat 240
189 acctttcatt caggaaaaat gtacttagac cctacaatgt actagtaggc ctctgcgctg 300
190 gcaatacaga taagataatg tagccccctg cctcaaagga actctcctcc ttaggttgca 360
191 tttgtataat gtttgatttt tagattgttc tttgagccct gcattccacg aggataggtc 420
192 agtgggtatt aacgaggtaa aaggggagta gtacgaaagg gcattcaagc gtcccattct 480
193 cgcttcaacc aaagcagccc tgcgttttcc tagttttatt aataggtttg atgtaaggtc 540
194 gtctttgaaa aggggggttg gctttttttt acagtgtgac tgaggataaa tttataaaaa 600
195 gggaaatgta tggcatggtg agttttttca catacatcct tgtgaatacc cagctcaaga 660
196 tccaaaacat tcccataatt tcagaaagtt ccaaaccctt gcctcttttc agtcttagcc 720
197 ctcttccctt gaagtaacca ctgttccgac ttcaatcact acttttatcc cacaggttaa 780
198 ttttttggtt tttttccact aaattttcaa attctttgat atggtacttt actattgacg 840
199 aagtactttc acactagggt atttaatat ctttgattca cccaatattt agggaacacc 900
200 tgtaggggac aaaaaatgaa tgagagcccc tgccttccat tgctgcta at ctggtgggaa 960
201 cgagacatgt atttaattaa gcatgtaaaa aatagagtgg gtgatgaa at aatctatata 1020
202 ctaaattccc atgacacaca gtttacctat gtaacaaacc tgcattgtga cccccaacc 1080
203 taaaatataa gttggaaatt aaaaaaaaaa gagagggaga atagagcatc acaaccagag 1140
204 tgctgagatg aattacttta ttaccaaaga aggaggagga ctacgggagg tgccgacgtt 1200
205 taaacccagt cactgaaggg tgtgcagaat ttggataggc aagataccct gggacaaggt 1260
206 cattctaaaa ccattgctaac atttgtactt tttttttcat tgtgatagtt cctgaaatga 1320
207 gttgcataaa actggtacat gtcttagggc agtctcta at tgatttttat tttgttctat 1380
208 ttttaaaaat tagtcttcaa atagcagatt cacatgat at taaaatatat gcacataaat 1440
209 tatatacaca aatataat tttt ctgaatgaaa ttttagtatc gcatatattt aagagctatt 1500
210 tetgtctcat atgttcataa tcttcaccca ttaaaaaaac ttttgtagg cctttctcac 1560
211 tetaagatta taaaaaat tccattatt tacctagcta gttttctagt tgttccaaaa 1620
212 ccatttattg aacaatccat ctttttgaca ctggtttggc atgcctta at tatatattct 1680
213 tgtgtgtggt aggatctcct tttggacttt ccattctgtt cattgagtct tatcagctcc 1740
214 tcttacattg gtaccatgat gttttaatct atggggcttt gtagtttaaa tgtagggcta 1800
215 gttccagcgc attgttctct atcagctggt aggaacttag aaatcagctt gctctgtttt 1860
216 aaagaaaaac ctggtatttt tttatcagta taacattcta tttatattaa cttgaagaat 1920
217 tgaaaacatc tatgattttt cctattcagt aacgtatcac ttagaatagg ttaggttgta 1980
218 ctactataaa atctcagctg cataaaacaa tttttttttg cttgtgctac acatccatta 2040
219 ggtcatcaag ggactcacct tgtcaagtta ctacagagatt caggctgata taaagggttg 2100
220 atcttgacat acgctttcat gatgacagaa agcaggggag agaagggtgt gagccatgtg 2160

```

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/10/775,501A

TIME: 11:47:24

Input Set : A:\84500-1.app

Output Set: N:\CRF4\09132004\J775501A.raw

```

221 ctttctcccc cttctatcca gaaatgacac atactcacat ttcattcgcc agagaaatta 2220
222 acatggcccc tcctaagttc aaatggatag agaaatgcct tcctaccagg tgcccagaat 2280
223 tagaagagca aacattttgtg aacagttctg agtaccacaa ataccgttat ctttccactt 2340
224 aagtcttctg tttcactcag tagtgcttta aacttttctt catatgtttt tcagtgtttc 2400
225 ttgttgaaatt tcttgatatt ttatcatggt tgttcgtact gggagtagcc tttttttcca 2460
226 tttcattttc tggctgggtt cattgctggt tgtttttttg ttttggtttg tttttgagat 2520
227 ggagtctcac tctgtcgccc aggtcggagt gcagtgtcac aatctcggt cactgcaacc 2580
228 tctgcctccc aggttcaagc gattcttctt tctcagcctc ctgagtagct gggattacag 2640
229 gcatgtgcca ccatgcccag ctaatttttt atatttttag tagagatggg gtttctccat 2700
230 gttggtcagg ctggtctcaa actcccaatc tcaggtgatc cgctgcctc tgccttccaa 2760
231 agtgctggga ttatagacat gagccaccgt gcctggccta gttcttatgg gatgtatatg 2820
232 tctttggatt catatgatat gtatatatgt ttatatttct acaagtacat acctaggagt 2880
233 ggaattgttg ggtcataggt taatgcatgt ttttctgcca aacagttgtg tcaatttctg 2940
234 ttttcaccgc tgtgaatgag agttgttcta ctttcttgac aacacttgat attgtcagtc 3000
235 atttttagcca ttctggtgaa tttatagtgc tatttctgtg tgtgtaagag agagaatgag 3060
236 agaggggtgtt tgtgagaaaa ccaaagcaac actgtgagag tgtgtgtgtt tgtgagaaaa 3120
237 ccaaaatata tactactgtg atttcattgg gagaaaatct gtttggtata tcaaaaaaag 3180
238 tagcttaatt acttcatcat tattggttta ggt 3213
241 <210> SEQ ID NO: 6
242 <211> LENGTH: 1296
243 <212> TYPE: DNA
244 <213> ORGANISM: Homo sapiens
246 <220> FEATURE:
247 <223> OTHER INFORMATION: lactase persistence type intron 9 of the MCM6 gene with
248 single nucleotide polymorphism (SNP) a substituted by g at
249 position -22018 5' from the intestinal lactase-phlorizine
250 hydrolase (LPH) gene
252 <400> SEQUENCE: 6
253 taagaacatt ttacactctt cagtataaag aagtcagaat acccctaccc tatcagtaaa 60
254 ggctataaag ttaccattaa aaagatgtcc ttaaaaaacag cattctcagc tgggcgcggt 120
255 ggctcacacc tttgtcccag tactttggga agccgaggtg ggtggatcac ctgaggtcag 180
256 gagttcgaga ccagcctggc caacatggcg aaaaccatt ttctctacta aaaatacaaa 240
257 aattagccgg gcattggtggc ggggtgcttgt ggtcccagct actcaagagg ctgaggtggg 300
258 aggatcactg agcccaggag gtggaggctg cattgagcca agattgtgcc actgcactcc 360
259 agcctgggtg acagagcgag actctgtctc aaaaaaacca aaacaaaaaa aaccagcat 420
260 tcttttagtaa ataattcata gttttcttca tctagaattt aaaattgtga tagttgatca 480
261 gcatgtcctg agcacgtgtg tttgctgtta ctagttaga tcggtagatg tgtatataag 540
262 ttataggtat aaaatcaatc ctgagttgac acaaggtttt gatgttgagt acaagtacag 600
263 taagtgtata tttttagtta tgccttagt ttttaagtcaa ttgtgtggtt ctttctagct 660
264 ttaggatctg ttgaattatc ttcttagaa aaggaggtta agaattctca cttacctatc 720
265 ttctacttgt ttggagaata gaagagtccc tgtggttagca gactttgtga gtttacttgt 780
266 aattttccat ctgaaagact gttcttgttt ttcgtgatga agtcttgctc tgtcgcccag 840
267 gctggagtgc agtgggtgcaa ccttggtcca ctgcaacctc tgcctcccgg gttcaagcaa 900
268 ttctcctgcc tcagcctccc gagtatctgg gattacaggt gcacaccacc acacctggct 960
269 aatttttcta ttttcagtag agacggggtt tcaccatgtt ggccaggctg gtctcgaact 1020
270 cttgacctca tgatcagccc acctcagcct tccaaagtgc tgggattaca ggtgtgagcc 1080
271 cccacactcg gccgttggtt ttttttaaga gacagggctc cactctgtca cctaacctgg 1140
272 agtacagtgg caatcatggc tcaactgtaac ctcaaagtc cggccttagt gaagcgttct 1200
273 tctgccttg gccctccaaa gtgctgggat tacaagtgtg agccatgcat ccagcttgaa 1260

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/775,501A

DATE: 09/13/2004
TIME: 11:47:25

Input Set : A:\84500-1.app
Output Set: N:\CRF4\09132004\J775501A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 1
Seq#:9; N Pos. 1
Seq#:11; N Pos. 1
Seq#:13; N Pos. 1
Seq#:15; N Pos. 1

VERIFICATION SUMMARY

DATE: 09/13/2004

PATENT APPLICATION: US/10/775,501A

TIME: 11:47:25

Input Set : A:\84500-1.app

Output Set: N:\CRF4\09132004\J775501A.raw

L:293 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
L:326 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:392 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:423 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0